DOCUMENT RESUME

ED 136 208 CS 003 258

AUTHOR	Weber, Jane E.
TITLE	The Kernel Distance Theory: Evaluation by Student Judgement of Sentence Difficulty.
PUB DATE	Jun 77
NOTE	115p.; M.Ed. Thesis, Rutgers, The State University of New Jersey
EDRS PRICE	MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS	Community Colleges; Junior Colleges; *Readability; Reading Ability; *Reading Research; *Sentence Structure; *Student Opinion; *Syntax; *Textbook Evaluation
IDENTIFIERS	De Pierro (Joseph); Fry Kernel Distance Theory

ABSTRACT

This study attempted to validate both the Fry-Kernel Distance Theory and the De Pierro findings, in determining the suitability of text material for the students to be instructed. First-year community college students were presented with sentences arranged in pairs, each representing two of four prescribed types. Subjects were to select the harder sentence of each pair. Sentence complexity in this study was measured through student judgment. Variables included reading ability, types of sentences, and student judgment. It was concluded that, although subject judgments on individual items may vary within sentence categories, when these were totaled, certain trends emerged: in two of the four types of sentence comparisons, a significant difference in difficulty was reported; for three out of the four hypotheses, good and poor readers did not differ significantly. However, further analysis indicated some tendency on the part of good readers to judge sentences in the direction opposite from that predicted. (FOD)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED136208

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE KERNEL DISTANCE THEORY - EVALUATION BY STUDENT JUDGMENT OF SENTENCE DIFFICULTY

A THESIS

SUBMITTED TO THE FACULTY

OF THE GRADUATE SCHOOL OF EDUCATION

OF

RUTGERS

THE STATE UNIVERSITY OF NEW JERSEY

BY

JANE E. WEBER

IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE

OF

MASTER OF EDUCATION

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Jane E. Weber
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

NEW BRUNSWICK, NEW JERSEY JUNE 1977

APPROVED: Edward Fry
EDWARD B. FRY, CHAIRPERSON

John Geyer
JOHN GEYER

Janet A. Emig
JANET A. EMIG

DEAN: ______________________
MILTON SCHWEBEL

CS 003 258

ACKNOWLEDGMENTS

The author would like to thank Dr. Edward B. Fry, committee chairperson, for his inspiration and guidance throughout the research and writing of this thesis. Appreciation is also expressed to the other committee members, Drs. John Geyer and Janet Emig for their cooperation and understanding.

Special appreciation is expressed to Dr. William Engler of Mercer County Community College for his cooperation arranging for the investigation, to his colleagues, Gertrude Hammerman, David Leeb, Walter Sikorsky, Nathan Albert and June Weinberg for allowing the use of their classes, and to the 105 summer school students for their help and enthusiasm taking part in the experiment.

The author wishes to acknowledge the contribution of Joseph De Pierro for his help in the preliminary stages of the research and to Mrs. Darlene O'Connell for typing the manuscript.

Special thanks are extended to Mrs. Gloria Lukacs, Reading Center secretary, and to Jonathan Weber, for their understanding and encouragement throughout the author's graduate student career at Rutgers University.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS ii

LIST OF TABLES vi

CHAPTER

I. INTRODUCTION 1

Background of the Problem 1

Statement of the Problem 3

Importance of the Study 4

Variables 5

Limitations of the Study 5

II. A REVIEW OF THE LITERATURE 6

Psycholinguistics and Reading 10

Syntax and Linguistic Theory 12

The Structural Approach 13

Finite State Grammars 13

Phrase Structure and Transformational-Generative Grammars 14

The Chunk Model 18

Semantic Impact on Language 18

Other Recent Approaches to Difficulty Identification 19

TABLE OF CONTENTS (continued)

CHAPTER Page

Background of Other Variables of this Study 26

Adult Population and Reading Ability 26

Subject Judgment 28

Summary of the Review of the Literature 32

III. PROCEDURE FOR THE EXPERIMENT 34

Samples 34

Sentence Production and Selection 36

Presentation 38

Student Responses 39

Statistical Evaluation 40

IV. RESULTS 43

Hypothesis 1 (D1 > D4) 43

Hypothesis 2 (D2 > D4) 45

Hypothesis 3 (D2 > D1) 45

Hypothesis 4 (D2 > D3) 48

Hypothesis 5: Reading Group Comparisons 48

Summary of the Results 51

V. DISCUSSIONS AND CONCLUSIONS 52

Comparison with Earlier Research 52

General Applications 57

Suggestions for Further Research 58

REFERENCES 61

TABLE OF CONTENTS (continued)

APPENDICES Page

A. Weber Sentences 67

B. De Pierro Sentences 72

C. Sample Questionnaire Form A 77

D. Sample Questionnaire Form B 87

E. Sample Answer Sheet 97

LIST OF TABLES

Table Page

1. Grade Scores of One ES100 Class On Test of Adult Basic Education - July 1976 Basic Composition - ES100 Section 9B24, MCCC . . 37

2. Tallies of Student Judgments of More Difficult Sentences H_o1: D1 > D4 44

3. Tallies of Student Judgments of More Difficult Sentences H_o2: D2 > D4 46

4. Tallies of Student Judgments of More Difficult Sentences H_o3: D2 > D1 47

5. Tallies of Student Judgments of More Difficult Sentences H_o4: D2 > D3 49

6. Comparisons of Good and Poor Reader Judgments of More Difficult Sentences, Hypotheses One, Two, Three and Four 50

7. Comparison of Results of Two Studies Based on the Kernel Distance Theory 53

CHAPTER I

INTRODUCTION

Background of the Problem

One of the major problems at all levels of education has been determining the suitability of text material for the students to be instructed. In order to assess the appropriateness of new writings, teachers have relied on three general approaches: (1) measuring the readability of the material with various formulas or scales, (2) testing the level of comprehension of a sample of students with queries on content from the text under consideration, and, most commonly, (3) using subjective judgment. Weaknesses exist for all three techniques.

The first readability formulas and scales to predict difficulty of written material were developed over fifty years ago, and today over forty exist. Most depend on a count of average sentence length, average syllables per word, an analysis of vocabulary and/or frequency of various parts of speech in a passage. Most of these do not directly consider the syntactic complexity of each sentence nor the background of individual readers for coping with various types of sentences. The major flaws of comprehension tests for prediction are that they are time-consuming to prepare and administer, may themselves suffer from problems of readability, and are only valid for the group of students and specific materials tested. The third approach is only as good as the experience of the teacher and his knowledge of his students.

During the past ten years a small group of investigators, influenced by recent developments in the fields of transformational grammar and psycholinguistics, have given more attention to the influence of syntax on readability. Attempts have been made to identify particularly easy or complicated sentence structures, and several new formulas and comprehension tests based on grammatical manipulation have appeared. Several new theories have also been proposed which should help eventually to improve predictive instruments. One of these is the Kernel Distance Theory suggested by Fry in 1974.

According to the Kernel Distance Theory, one method by which sentence complexity may be altered is by varying the placement of subject, verb, and object in relation to each other, and to the sentence as a whole. Four basic patterns of varying complexity (and, consequently, of reading ease) are postulated. Vocabulary, sentence length and any other factors usually measured by traditional readability formulas remain constant across all the sentences compared.

Thus far one study to validate the Kernel Distance Theory has been carried out. The Joseph De Pierro dissertation at Rutgers University has developed a set of sentence pairs, each of which contrasts two of the four basic patterns. In order to determine differing sentence complexity De Pierro measured subjects' silent reading time while they read each sentence of a pair, and their degree of accurate recall and response delay of each.

Statement of the Problem

This study was an attempt to validate both the Fry-Kernel Distance Theory and the De Pierro findings. In this study, two samples were utilized from one grade level, the first year of a community college. As in the earlier investigation, subjects were presented with sentences arranged in pairs, each contrasting different patterns though made up of the same words. However, in this study the same sentences and some new ones were presented in a classroom setting, and sentence complexity was measured through student judgment rather than the dependent variables of the earlier study. Students were asked to report the harder sentence in each pair and to indicate their selection by checking the appropriate blank on their answer sheets. Results were analyzed to identify more difficult sentence patterns and to see if subjects of different reading ability react in different ways. Hypotheses to be tested:

1. Subjects will judge sentences that have distance at the beginning harder than those that have distance at the end (D1 > D4).[1]
2. Subjects will judge sentences that have distance between subject and verb harder than sentences with distance at the end (D2 > D4).
3. Subjects will judge sentences that have distance between subject and verb harder than sentences with distance at the beginning (D2 > D1).

---

[1] According to the notation system used here, D = distance (such as a clause), and numerals indicate where that distance occurs in the sentence: 1 = before the kernel, 2 = between subject and verb, 3 = between verb and object, and 4 = after the kernel. The word "kernel" as used in this study always refers to the subject, verb, object combination, the core of a sentence.

4. Subjects will judge sentences that have distance between subject and verb harder than sentences with distance between verb and object (D2 > D3).
5. Subjects with higher reading scores will judge sentences differently from subjects with low reading scores.

Importance of the Study

It is true that traditionally investigators have shied away from investigations based on subjective evaluation. A few studies exist which utilized student evaluation of specific text material or judgment of word frequencies. Otherwise most opinion measurement surveyed in the general literature appears to have been confined to evaluation of attitudes toward social institutions or activities. However, it is the belief of this researcher that student judgment of text, although subjective, is a legitimate area for research. After all, if a student thinks a sentence is too difficult to understand, though he may be interested in the subject matter, he may decide not to continue reading a passage. He will no doubt not understand why he finds it so difficult if the problem is syntactic, but for him that is immaterial. It is not an immaterial concern for the writer or text evaluator however. If he can identify various syntactic structures that a population of subjects with a predetermined reading ability and scholastic background thinks is too complicated for reading comprehension, he may be able to produce or select more appropriate reading materials for his reading audience.

Variables

1. Reading ability: This was determined by examination of scores on reading achievement tests. For this study, scores were taken from the reading portion of the College Guidance and Placement Test. Cut-off points used by the subject school to determine "good" and "poor" readers were utilized to categorize students accordingly.
2. Types of sentences: These followed the four basic patterns suggested by Fry in the Kernal Distance Theory and described in the first four hypotheses above.
3. Student judgment: Subjects were asked to choose the more difficult of two sentences. It is hoped that accurate student responses were encouraged through carefully designed instructions and the oral introduction.

Limitations of the Study

The investigation was restricted to one college grade level, although a span of age was expected. Testing was administered in a general classroom setting and only once to each subject. Consequently, it was not possible to take an average of individual performance. The only sentence patterns to be manipulated were those indicated above, and applications of results will, therefore, only be applicable to such sentence types. As in the De Pierro study, sentences were presented in isolation and findings may not be applicable to student judgment of such sentences in context. Perhaps the biggest limitation of the study was the reliability of subject responses. Finally, no attempt was made to apply any findings to existing readability formulas or to develop a new one.

CHAPTER II

A REVIEW OF THE LITERATURE

This review of the literature will begin with a brief consideration of the development of readability analysis, with an emphasis on its history and traditional weaknesses. In order to understand these weaknesses, this chapter will focus next on recent models of reading developed by psychologists and psycholinguists, and on recent theories of linguists, which, in turn, have affected the models of the reading process. Research employing the specific variables of this particular study will then be analyzed--syntactic manipulation of sentences, use of college populations for data collection, differences between good and poor readers, and subject judgment as a measuring device.

Readability research has been going on since shortly after World War I. Initially it was primarily concerned with tailoring instructional materials to students and helping to determine the appropriateness of commercially available materials. In the course of study such research has also helped provide insights into the nature of the comprehension process.

Without systematic means of predicting student success with particular materials, the average teacher to this day has two means of making such determination. He or she can use a test for judging comprehension difficulty, or can rely on his or her previous experience with such materials and groups of students. The former solution introduces two problems: such a test is typically in itself, a reading task with which some students may not be able to cope for various reasons, and it does not help pinpoint specific problem areas, i.e., words, phrases and/or sentences. Teacher judgment has been shown to be impressive in many cases (Jorgenson, 1975). However, this technique appears to be tied to teacher expectations for their students, and could hardly be considered acceptable procedure for the inexperienced. Another problem is that as with the testing approach mentioned above, no means for close analysis of potential problem areas is provided, nor is guidance given to future writers or revision editors. Hence, the need for a more scientific approach.

Readability formulas originated with that of Lively and Pressey (1923) and by the time Jeanne Chall wrote her exhaustive appraisal of those existing by 1958, twenty-nine separate formulas or similar measures of readability could be evaluated. Of these, the majority relied on analysis of vocabulary, comparing samples of text to basic word lists, such as Thorndike's, published in 1921. Other criteria variously considered included a number of different words used in a passage (as in Washburne and Vogel, 1928), words beginning with crucial letters (Lewerenz, 1929), average sentence length (a common measure, first used by Gray and Leary, 1935), and such abstract measures as "human interest" or "level of abstraction" made popular by Flesch (1948, 1951). In a few cases, consideration was made of syntactic elements in the selections, which we shall see was of prime interest in the current study. Of these, the most noteworthy were Washburne and Vogel (1928) who considered numbers of prepositions, Ojemann (1934) who employed counts of prepositions plus infinitives, Dale and Tyler (1934) who utilized the number of indeterminate clauses, Gray and Leary (1935) who counted prepositional phrases, as did Lorge (1939) in his first readability formula. However, attempts to systematically assess the validity or reliability of such measures were limited at best.

In subsequent years, readability formulas tended to eliminate direct syntactic measures. The trend was toward achieving maximum objectivity, simplicity and ease of computation so that any clerk, or eventually, any computer, could be trained to quickly and accurately predict the readability of a given passage. So, we note that by 1963, when Klare was writing his review of the field, most practitioners were relying on formulas which correlated average sentence length, considered to be a sufficient measure of syntactic complexity, with some measure of vocabulary, such as a comparison with a basic list, or merely an average of syllables per passage. Of course this incorrectly assumes that all words of equal length, and all sentences of equal length, are equally difficult. However, such techniques were still being staunchly defended in 1974 by Glazer whose studies showed that most longer sentences also happened to be comparatively syntactically complex.

In 1974, Klare reviewed the latest advances in readability formulas. Of thirty English formulas and methods examined, only seven took syntactic complexity into account. In that same year, Harris noted that from the research he surveyed the most consistently reliable variables of readability were vocabulary difficulty and sentence length. However, he placed much credence in current research into the effects of various linguistic variables, which will be described later.

According to Bormuth (1967) two important developments have taken place in the area of readability in the past twenty-five years. One is the development of the cloze technique (Taylor, 1953). Briefly, this method of analyzing text requires that words be deleted according to a set schedule, usually every fifth word; difficulty is then assessed according to the level of performance of subjects at restoring the missing words. Bormuth and others claim this method to be highly valid and reliable. They also point out that it helps isolate passages of especial difficulty.

However, other researchers question these findings. Hittleman (1973) pointed out that the deletion of certain types of words, i.e., content words, can be more deleterious to performance than removing structure words. Wood and Schlieff (1974) also found little correlation between cloze measures and those of nine prestigeous readability formulas for determining the difficulty of non-text materials. Lastly, Carver (1973) found from studies with college level subjects that, though cloze may be appropriate for testing the recall of memorized materials, this technique is not valid otherwise for testing comprehension. He implied that this technique is not appropriate for assessing readability for this population as well.

Bormuth pointed out one other important development in recent years: the publication of two comprehensive reviews and analyses of the field of readability, one by Jeanne Chall in 1958, and the other by George Klare in 1963. These have helped point the way towards the development of more reliable methods of measuring reading difficulty and better analysis of obtained data. Both also emphasized the need for accurate measuring and describing of linguistic features which affect comprehension.

A third major development, not mentioned by Bormuth should also be noted. In 1962, Ruth Strickland published the first major study by an educational researcher of the language complexity of children. Her prime concern was the influence of oral language patterns on reading comprehension. In her studies, she concentrated on oral language units, i.e. those separated by intonation. This paved the way for others to consider language units separate from literal sentences.

In order to better understand the change in emphasis from simplistic sentence element counts to the complex linguistic analysis on which the current research was based, it is important to consider the nature of the comprehension process in reading, as well as linguistic features of language.

Psycholinguistics and Reading

Perhaps the first psychologist to seriously consider the influence of syntax on thought, thereby kicking off the whole new field of psycholinguistics, was George A. Miller. In his presidential address before the Eastern Psychological Association in 1962, he challenged his cohorts to accept a more realistic conception of language. Generally speaking, psycholinguists claim that readability can only be understood within the context of the reading process itself. In order to read, the reader utilizes his linguistic and experiential background, along with the graphic display on a page in order to process three kinds of information: syntactic, semantic and grapho-phonemic. Reading then becomes a matter of using a variety of cues to make tentative decisions which are confirmed or rejected as reading progresses. This viewpoint was first presented by Thorndike (1917) and has probably been best developed recently by Goodman (1970).

Other psycholinguists appear to agree. Smith (1973) divided reading into two concurrent processes, one utilizing visual information, and the other, nonvisual information, and he suggested that there may be a trade-off between the two. If one is readily familiar with the language and content of a passage beforehand, the reader may find material easier since less visual information is needed. The converse is also true.

Several experiments which have analyzed reading errors (or "miscues" as Goodman would call them) appear to bear this out. They have shown that most skilled readers do not attack a passage word by word, but in grammatically acceptable sequences. When errors are made, they are most usually substitutions by words which are the same part of speech as the word misread (see Kolers, 1973 and Hoffner, 1974). This tendency to make syntactically acceptable errors appears to be typical of even low-ability readers, although they appear to be less able to retain long units of cues in memory.

Hoffner concluded that syntax was most important for cueing in the easier passages; in harder ones, readers fell back on the graphophonemic cues, although the syntactic cues still appeared to be more important than the semantic. In other words, she found that "the ability to process syntactic features did not depend on the ability to process meaning" (p. 85).

Hittleman (1973) has summarized very nicely what these findings mean for readability measures. He suggested that because individual readers' backgrounds are not taken into account, the traditional formulas, based as they are solely on the printed page, can in no way be considered absolute measures of readability. Considering this problem, Bormuth has suggested that the only way out may be developing readability profiles. These would include levels of difficulty of various language features in a passage as determined for various groups, rather than the constant correlation between features which typify formulas now in use. Study of linguistic variables, the major activity of much psychological research of late and of this thesis, should be particularly useful for developing such profiles.

Syntax and Linguistic Theory

The evolution of linguistic theory has been a rather complex process which has snow-balled in the past quarter decade. Several important new schools of thought have emerged which have augmented, and in some cases, contradict, earlier held beliefs, as well as each other. Since many detailed texts have appeared on the subject, only a cursory description will be given of the major theories currently supported, with an emphasis on syntactic aspects of each. Greater emphasis will be given to recent psychological experiments designed to verify these theories.

The Structural Approach

Up until the mid-1950's, most of linguistics was descriptive, or structural as it was often called. Such an approach seeks to identify sound features that identify and separate words, markers that identify units of grammatical patterns, and patterns of grammatical structure that "regularly elicit recognition responses of grammatical or structural meanings." (Fries, 1963, p. 73)

This approach is often contrastive and of especial interest to foreign language teachers. For instance, in modern English it is essential to understand the position of the subject and object in each sentence because there are no contrastive forms of the article "the." In the sentence, "The man killed the lion," reversing the position of the subject and the object would produce an entirely different meaning from the original. This is not true in many other languages.

Interesting though this kind of analysis is to the study of language, it does not lend itself well to experimentation nor to explanation of language behavior. It can be used for language exploration, however, as in the studies of Harrell (1957), Strickland (1962), and Loban (1963), which all analyzed children's language. Others have demonstrated that employing such oral language patterns in reading passages can lead to higher comprehension scores for the analyzed subjects (see Ruddell, 1963 and Tatham, 1969).

Finite State Grammars

Behavior theorists during the 1950's postulated that language is produced in Markovian chains, each nexis of which limits possible options for the next state of the system (Hulse, Deese and Egeth, 1975).

This means that the selection of which word follows another in a sentence is determined and limited by previous words. If the first word chosen is the English article "the," the next word is confined to all English nouns, adjectives and quantifiers; verbs and conjunctions are not permitted according to the rules of English.

More recently, this theory has been rejected by most psychologists and linguists, although it can be used to artificially generate acceptable sentences. It has been found that certain kinds of grammatical constructions cannot be generated by finite state rules. Perhaps more important, finite state grammars simply do not agree with what we know about sentence production. Usually we do not generate a word at a time; frequently, we have an idea of the whole sentence before it is expressed, the end of which often determines the beginning.

Phrase Structure and Transformational-Generative Grammars

Perhaps the most important development in the field of linguistics in recent years has been the concept of phrase structure grammars, first devised by Zellig Harris and Noam Chomsky (1957, 1965). They proposed that there exists within each individual a set of inductively formulated rules by which only acceptable sentences are generated. These sentences, rather than being described as chains in this system, are usually portrayed as trees, which branch downward. Phrase structure grammars begin by dividing sentences into phrases, which are further divided, depending on their complexity.

The most important aspect of their theory is the distinction between deep and surface structure. Deep structure is the level at which semantic rules operate, which allows the reader/listener to understand the meaning of a sentence. Some theorists believe deep structure is composed of simple active sentences (sometimes called kernel sentences) and are the input of memory. Surface structure, on the other hand, is the language to which the individual is actually exposed, which is mapped out by phonological rules. The two levels are linked by explicit syntactic rules which generate various complex structures from the simple kernel sentences. In other words, all sentences of a language are either kernel sentences or transformations of these kernel sentences (Fries, 1963). It should be noted that the notion of kernel sentences is usually no longer used; most researchers tend to refer to the content of deep structure as an abstract structure. However, kernel sentences are sometimes assumed to resemble deep structures more closely than any other type of sentence (Howe, 1970).[2]

Psychologists have maintained that a sentence becomes more complex depending on its inherent transformations. As more transformations are added or different kinds are required, this will be reflected in the passage complexity and subsequent subject comprehension. Countless experiments have been conducted to verify this theory and, to aid readability, to establish a hierarchy of transformational complexity. Some of the more prominent experiments in this field are mentioned below.

Perhaps the first experiments in this area were performed under the leadership of G. A. Miller, as reported in 1962. He described a series of studies in which subjects were presented with series of sentences produced from kernels--negatives, passives, and passive-negative transformations. Findings gave broad, general support to the transformationists with some qualifications. Mehler (1963), using a similar design, found that most subjects could recall sentences of accurate meaning more easily than those with precise grammatical constructions.

---

[2] It is of great importance to note that a kernel sentence is not the same as a sentence kernel. The Kernel Distance Theory tested in this thesis uses the term operationally to refer to the core (subject, verb and object) of any surface structure, rather than in the sense described above.

E. B. Coleman (1964, 1965) compared subject memorization of four types of grammatical transformations: nominalizations, actives versus passives, adjectivizations, and embedded versus nonembedded sentences.[3] He found that active verbs and actives, and nonembedded sentences were easier to learn than their counterparts, but found no significant difference between adjectivizations and their counterparts using adjectives. The nature of the sentences stored in memory was also explored. It was determined that they were probably not kernel sentences. However, he also reported a tendency of subjects to recode passives into actives.

Gough (1965) researched the effect of different grammatical transformations on subjects' speed of understanding. Active sentences, affirmatives and true sentences were verified faster than their counterparts. The author did indicate, however, that more than syntactic complexity was being tested, as indicated by the true-false variable. Similar effects were found when free recall (Howe, 1970) and ease at a deductive reasoning task (Lippman, 1972) were used as measures of psychological complexity when varied transformations were presented.

---

[3] Embedded sentences will be referred to later in this thesis as those with subject-verb or verb-object splits; nonembedded sentences are those with distance at the beginning or the end instead of within the sentence kernel.

Another phrase structure approach was developed by V. H. Yngve (1960). Like the Chomskian method, it begins with a tree diagram branching downward and its complexity is determined by the complexity of the sentence. However, it also assumes that grammatical structures are produced from left to right. Only one part of the structure can be completed at a time and, while doing so, other parts must be held in memory. This model assumes that the difficulty of the structure of a sentence derives from the number of grammatical constituents that must be stored in memory as a sentence is produced or interpreted. Also, therefore, the longer the sentence, the harder. To determine relative difficulty, a set of counting procedures is applied.

Several experiments exist based on the Yngve model and its predictions, and a few also compare these with those of the Chomsky model, whose predictions are often diametrically opposed. Mehler and Carey (1968) studying student accuracy with sentences containing transitive verbs and predicate nominatives, found support for the Chomskian approach while Rohrman's experiments to investigate whether surface or deep structure is more crucial in storing and remembering sentences obtained results supporting the Yngve model (1968).

Perfetti (1969) also attempted to compare the sentence depth measures with transformational complexity using an experiment requiring recall of sentences. In this case, results indicated no support for either approach. It should be pointed out that some researchers do not consider the depth approach adequate to describe all sentences. More than one depth count is also possible for some sentences, which can be represented by more than one diagrammatic structure (MacGinitie and Tretiak, 1971).

The Chunk Model

The latest model to emerge in the literature is the chunk model, described by Bransford and Franks (1971). This is a completely different approach from the phrase structure analyses since it emphasizes improvement of comprehension with the generation of larger conceptual or semantic units. That which the chunk model predicts will be difficult, the earlier models (especially the deep structure model) says will be easy. It should also be noted that the chunk model conflicts with readability formulas which equate sentence length and grammatical complexity with more difficult comprehension, as we have seen. One study has been found comparing the various approaches (Pearson, 1975). Here, grammatical complexity was found to be an aid to comprehension and recall in many cases rather than a hindrance, thereby supporting the chunk model. Perhaps, Pearson suggests too much attention has been paid to syntactics and not enough to other aspects of language.

Semantic Impact on Language

Research on linguistic variables has centered primarily on syntactic manipulation, perhaps because of the presence of several strong theories in this area. Some researchers, however, have taken a closer look at the influence of semantic factors. Actually, it is included as one aspect of the Chomskian model, the deep structure, but has been relatively ignored compared to the transformational aspects. Aside from Pearson, several other researchers have given some specific attention to the effect of meaning on the reader in different grammatical contexts. Gough (1965), Mehler and Carey (1968), and Slobin (1966) have all suggested that in certain grammatical structures, such as negative sentences or in tests of veracity with varied grammatical presentations, the semantic element has strong impact on subject responses. Most recently, Hansell (1976) has concluded from the ambiguous results he found in a series of syntactic and semantic mutilation experiments that the deeper causes of language difficulty have yet to be identified.

Other Recent Approaches to Difficulty Identification

Since virtually all linguists, psychologists and readability experts acknowledged that the causes of language difficulty have not yet been isolated, the search for new approaches and answers continues, some along older lines of procedure, a few along new ones.

Kellogg Hunt has returned to a structural linguistic approach, identifying language maturity and areas of potential problems by determining "minimal terminable" or "T units" through examination of students' writing (Hunt, 1970). John Spencer has suggested the importance of what he calls "collocation," the tendency of certain items in a language to occur close to each other, and "set," the tendency for such word groups to overlap (described in Moir, 1970). Ruddell has pointed out that one area frequently left out of psycholinguistic considerations is the individual's interests, attitudes, and values. He has given these affective factors equal weight with cognitive strategies in his communication model (Ruddell, 1970).

Two attempts have been made to create measures of syntax in reading. Marcus (1971) has designed a diagnostic test to identify weaknesses in seventeen specific skills. Using items developed from kernel sentences according to transformation-generative theory, this instrument examines a wide range of structures, as defined by the linguist, W. Nelson Francis. The author claims high internal consistency based on its use with 421, fifth through eighth grade subjects.

In 1972, Botel and Granowsky published a Syntactic Complexity Formula, also based on transformational-generative grammar theory. Various sentence patterns were assigned different weightings, depending on their relative complexity. The authors suggest that this formula be used along with a vocabulary predictor in order to evaluate a wider range of linguistic variables than the syntactic alone.

Reaction to this formula has been mixed--most evaluators suggesting that it is a good beginning, but needs further refinement. Others consider it too simplistic, too arbitrary in the assignment of weightings, or too time-consuming, considering the crudeness of the instrument.

In 1974, Edward Fry proposed another approach for assessing passage difficulty. Traditional measures of difficulty, as we have seen, rely on counts of sentence length and average word length in syllables or word frequency. Fry has pointed out that this is not sufficient since sentences may be increased in difficulty through manipulation of word order, even though sentence length and vocabulary are held constant. He attributes the varied difficulty to the location of, and distance between, subject, verb and, when present, of object.

Fry defines these three elements of a sentence, subject, verb, and object, as the sentence kernel. As previously noted, this is a different concept from that of "kernel sentence" as used by transformational theorists. Simply stated, the resultant Kernel Distance Theory postulates that the nearer the kernel is to the beginning of the sentence, and the less distance between the elements of the kernel, the easier the sentence. The theory also states that distance between the verb and the object may create less difficulty than distance between the subject and verb.

This is not the first time researchers have considered the impact of subject-predicate relationships on sentence comprehension. In the Yngve model mentioned above, it was noted that while one part of a sentence is developed, e.g. the predicate, the rest of the sentence must be held in short term memory. If a complex subject were included in a sentence (referred to often as a left-embedded sentence), so much more information would have to be remembered before one moved on to the verb. Also, as noted above, structural linguists have long recognized the importance of subject-predicate position in English sentences.

Some researchers have experimented specifically with subject-predicate and object relationships. Herriot (1968), requiring substitution of English words for nonsense words in grammatically structured nonsense sentences, found that subject and predicate are indeed important components of sentence processing. His experimental group expected the grammatical subject to occur first in a sentence, but indicated more uncertainty with the verb and object. Fodor, Garrett and Bever (1968) tried to analyze verb complexity. They found the semantic component of the anagram tasks they required to be essential for syntax recognition. They also found evidence of assymmetries of lexical requirements in English sentences between the parts of the sentence on each side of the main verb.

In 1971, Hamilton and Deese reported on an investigation into the influence of the surface form of complex sentences and the relationship between subjects and verbs. They found that comprehensibility of sentences depended very heavily on the contiguity of the main elements of each phrase, irrespective of length. They also found that center-embedded sentences (ones with clauses separating subject from predicate) were much harder for their experimental subjects to comprehend.

It should also be noted that the syntactic measurement techniques of Marcus, and Botel and Granowsky described above also considered placement of clauses in relation to subject, verb, and object. After some experimentation with his diagnostic test, Marcus concluded that, "complex sentences in which a relative clause interrupted the subject-verb-object sequence of the independent clause was more difficult for the students to understand than complex sentences whose basic components were not separated." (Marcus, 1971, p. 58)

Fry's theory has suggested that this dimension could be added to current readability formulas while increasing their accuracy, especially with short passages. It also may have implications for writers who are concerned with lowering the readability of their material. However, before this is done, there is a need for further verification of the theory. The first experiment based directly on the Kernel Distance Theory has been conducted by Joseph De Pierro (1976) at Rutgers University. Since this experiment is the basis for this researcher's work, it will be described at length.

De Pierro translated the Kernel Distance Theory into four hypotheses:

1. When linguistic distance is preposed, the sentence will be more difficult than when linguistic distance is postposed.

   Example:

   Preposed - After eating supper, I am no longer hungry.

   Postposed - I am no longer hungry after eating supper.

2. When linguistic distance splits the subject-verb link of the sentence, the sentence will be more difficult than when linguistic distance is postposed.

   Example:

   S-V Split - The city, after ten years of fighting, fell in flames.

   Postposed - The city fell in flames after ten years of fighting.

3. When linguistic distance splits the subject-verb link of the sentence, the sentence will be more difficult than when linguistic distance is preposed.

   Example:

   S-V Split - Ted, nearly dropping one of his crutches, whirled around and cried.

   Preposed - Nearly dropping one of his crutches, Ted whirled around and cried.

4. When linguistic distance splits the subject-verb link of the sentence, the sentence will be more difficult than when linguistic distance splits the verb-object link.

   Example:

   S-V Split - The railroad, a few months after the bad accident, knocked down the shaky bridge.

   V-O Split - The railroad knocked down, a few months after the bad accident, the shaky bridge.

These hypotheses were tested with two samples: a college age group and an upper elementary level group. Each subject was presented with sentences which originated from materials appropriate to each level. For the college level sample, the basic sentences were selected from articles, stories, and paragraphs contained within several college reading textbooks: The Improvement of College Reading, Power in Reading Skills, The Meaning in Reading, Read with Speed and Precision, and Improving College Reading. Using basic sentences from such sources was considered to be more valid than fabricating sentences, which might have been easier. Sentences were then manipulated to provide appropriate linguistic distance and to facilitate experimental procedures. They were then arranged in pairs and presented to subjects individually on index cards.

Four pairs were presented for each hypothesis. Each subject was required to deal with 16 pairs or 32 sentences in all. So that no syntactic variation appeared in any position a predominant number of times, items were presented in a scrambled fashion. Subjects were told to read each sentence once, to remove the card and then to repeat out loud as much as they remembered. Four dependent variables were measured and analyzed: words recalled, silent reading time, combinatory quotient, and response delay. Results were subjected to a wilcoxon matched-pairs signed-ranks test.

De Pierro found that in some cases subject-verb splits were more difficult than when they remained intact, but he discovered little evidence of problems with verb-object splits. Difficulties with preposed and postposed patterns occurred for the college sample, but not in the direction predicted. Since De Pierro's samples were small and not controlled for reading ability, and since his results did not follow predictions, further research along these lines is warranted.

It should be pointed out that De Pierro's dependent variables also have their weaknesses. Recall techniques, silent reading time, and response delay are all indirect methods of assessing processing and comprehension difficulty, and may be affected by intervening variables. The time measures particularly may not be valid because of individual reading styles or inadequacies; they may, therefore, be measuring reading ability rather than readability, especially with the first presentation of a sentence. Recall, too, may not actually measure readability and may be heavily influenced by storage capability. Experiential background or possibilities for imagery production could allow differential recall of various sentences.

De Pierro acknowledges other problems with response delay measures. The manual measures he used proved to be inaccurate and difficult to employ. Some subjects were also found to reiterate a sentence as quickly as possible regardless of how accurately they performed, whereas the response style of others was more relaxed.

One must also examine De Pierro's subject groups. Two completely different samples were utilized, one relatively proficient and adult, and the other, still in the process of acquiring language as well as reading behavior. A wide range of ability was also found to be present within both groups on previously administered reading tests. In sum, too many variables might have been present to draw conclusions about each sample and differences between them.

Background of Other Variables of This Study

Adult Population and Reading Ability.--

In this research, the sample was restricted to a community college group alone. Such students have been of interest of late to readability experts since many of them are apparently ill-prepared for college level reading (Cline, 1972-1973). However, research with adult groups, in general, has been around since the early attempts to establish readability formulas.

According to Klare and Buck (1954) the first measure appropriate to adult materials was Kitson's, developed in 1921. This was a comparison method, not a true formula. The first formula for adult materials was the Dale-Tyler, developed thirteen years later, and the exhaustive readability studies of Gray and Leary were based almost entirely on studies of adults' reading habits. By 1954, when Klare and Buck's Know Your Reader was published, fourteen methods for assessing adult materials had been developed; four additional ones were accounted for by 1963 (Klare, 1963), and at least four more have been developed since then specifically for adults (Klare, 1974).

In addition, the cloze technique has been used with varying amounts of success with adult subjects and materials.

Of the linguistic studies mentioned earlier in this chapter, many employed undergraduate or adult populations. These included those of Bransford and Franks (1971), Mehler (1963), Perfetti (1969), Lippman (1972), Fodor, Garrett and Bever (1967) and, of course, the De Pierro study (1976).

With the emergence and growing importance of community colleges in the past ten years, several studies have emerged to assess their special problems. Ohaver (1971) conducted a study to examine semantic and syntactic cueing by low reading ability college students and found they had special trouble using semantic cues. It is interesting that these results concur with those found by Guthrie (1973), who used younger readers in a similar experiment. Here, too, disabled readers were found to be deficient in amount of comprehension, but not significantly different from normal readers at syntactic processing during silent reading.

Other research has considered the appropriateness of assigned texts in community/junior colleges. Hagstrom (1973) evaluated twelve assigned texts using the Dale-Chall formula and found that nine were more than one year above the reading level of two-thirds of the students, whose ability was determined by the Diagnostic Reading Test, Form A. Cline (1973) used the Nelson Denny test scores of close to 300 subjects and matched them against the Dale-Chall readability levels of seventeen texts. This time, eleven were found to be above the reading level of fifty percent of the students.

Spring (1975) used the cloze procedure to measure student success with six texts, which were of varying levels of difficulty according to the Fry nomograph. Twenty-three percent of the students were found to be reading at the frustration level, and fifty-seven percent at the independent level. However, Spring also discovered little correlation between reading ability and student grades. Apparently, many who, according to the cloze test, cannot read their texts, are getting a large amount of information elsewhere.

Subject Judgment.--

Experimenters have used a wide variety of variables to assess subject processing of different forms of syntactic and semantic constructions. Among those mentioned in the research cited above were word or sentence recall, time measures such as silent reading time and length of response delay, eye movement and eye-voice span, performance on cloze tests, along with traditional measures of comprehension using written questions. One under-utilized assessment technique is the subjective, i.e., asking subjects for a judgment, preference, evaluation, or their general reaction or attitude towards a discrimination task or such a task's components.

Most of these factors have not been adequately defined and researched in current psychological literature. They are usually subsumed under attitude theory.

Attitude has been defined as, "a complex psychological construct, built from the theoretically subordinate constructs, habit, cognition and emotion." (Greenwald, 1968, p. 386) It is often thought of as the result of classical conditioning, but some psychologists also acknowledge some genetic influence. Attitude theory, a very large and complex branch of psychology, made great strides during the period from 1930 to 1950, particularly with the development of standard procedures for measurement by Thurstone, Likert and Guttman, and continues to be of great interest to many today.

Of the recent general texts published on psychological measurement, virtually all define attitude as feelings in relation to a class of objects, ideas, institutions or people, almost always in a social context. Scales and other measuring devices have been designed accordingly. The most important are those of Thurstone and Likert.

The Thurstone scale is often referred to as a consensual scale. It is developed by having independent judges rank statements of attitude about an institution, e.g., the church. Once position of each statement is established, they are scrambled and presented to subjects. Subjects are then asked to either agree or disagree with each statement. Their position on such a scale is determined by calculating either the mean or median position of the items they accept. As should be evident, this scale is most appropriate for assessing attitude when a wide range of opinion may be held about objects or institutions under study.

The Likert, or summative scale, does not take position of items into account. Here, the subject is presented with a statement and asked to indicate his extent of agreement or disagreement along a continuum, the central choice of which is usually "undecided." Usually about an equal number of opinion statements favorable and unfavorable to the issue are presented. Attitude is assessed by counting the number of favorable or unfavorable statements the subject is willing to endorse. It should be noted that neither the Thurstone, nor the Likert scale may necessarily be appropriate for evaluation of student judgment of different grammatical configurations.

More recently, psychophysicists led by S. S. Stevens of Harvard University have refined a method of magnitude estimation (Stevens, 1956). In his experiments, subjects were presented with auditory stimuli and asked to specify the ratio of loudness between them. The magnitude estimations obtained through such subjective judgments were consistent with a physical scale of loudness. Subsequently, several psychologists have adapted this technique for other types of research, as will be described below.

Attitude assessment and subjective judgment have been used in educational research, in readability studies, as well as in studies of syntactic and semantic manipulation. The first quantitative study in readability, by Lively and Pressey (1923), utilized examiner scaling of eleven books and one newspaper, which was then used as the criterion for judging other materials. Washburne and Vogel (1928) were the first to rate reading interests based on an impressive survey of thirty-seven thousand children in the Winnetka, Illinois school system.

Reader reporting was a major tool utilized by Gray and Leary to analyze adult reading habits (1935). Not only were respondents asked what books they had read, but also what factors they felt added or detracted from the book's readability. Influential factors were then classified where possible into four categories: style, content, format, and organization. These, in turn, became the basis for further evaluation and research. A similar survey of high school and college students was subsequently conducted by Ruth Strang (1938). More recently, Hackman and Kershner of the University of Maryland used reader judgment along with length of reading time as criteria of readability (as reported in Chall, 1958).

In 1971, Estes developed a scale to measure attitudes towards reading to further teachers' understanding of their students. A pool of statements was collected from a group of twenty-seven teachers from one school district, and a Likert scale was developed for use with a large heterogeneous group. An assessment of general student attitude was made along with an item analysis. Many surveys of student reaction to courses, teachers, and texts, including a few mentioned earlier, have also been located.

As reported by Shapiro (1969) and Carroll (1971), a number of investigators have demonstrated that subjective scaling of relative word frequency may be an alternative to objective counting techniques. Both of these researchers attempted to verify this using Stevens' subjective magnitude estimation method with a wide variety of groups. In virtually all cases, averaged data correlated highly (over .90) with objective frequency counts. Deviations, Carroll noted, may, in fact, be a result of more valid subjective estimations, since established counts such as the Thorndike-Lorge may be dated and are subject to sampling biases, especially in the case of words of low probability.

Research of linguistic variables has also relied more of late on student judgment. The Hamilton and Deese study (1971) on comprehensibility and subject-verb relations is often used as a model. In this experiment, subjects were asked to indicate whether a sentence was comprehensible and to rate their confidence in that judgment on a seven point scale. A similar approach was used by Hansell (1976) to assess student reaction to passages, although the scale was reduced from seven points to five.

Lippman's syllogistic reasoning task (1972) used a different but equally common technique, requiring subjects to judge each problem on an eleven point scale for difficulty after they had completed each solution. Pearson (1975) used a third approach. In one series of experiments, subjects were asked to rank four sentences from the best, easiest, and clearest (assigned a rank of 1) to the worst, hardest, and least clear (to be rated a 4).

Summary of the Review of the Literature

Research into the nature of readability now spans a period of over fifty years. It concentrated first on identifying easily replicable counts of sentence components to be used for predicting passage difficulty. Advances in the fields of linguistics and psychology, however, have led to a reassessment of established techniques. Psycholinguistic models of comprehension, as well as linguistic theories of language production and understanding, have laid the basis for vast quantities of controlled research on many variables of language. More accurate measurement techniques and better understanding of the mathematics employed by the various formulas has led to refinements of earlier instruments.

However, the research points out that educators are still a long way from having a tool to easily and accurately assess reading materials. Many varied studies of linguistic variables--syntactic, semantic and phonologic, using all types of populations and measuring devices, are needed as yet.

CHAPTER III

PROCEDURE FOR THE EXPERIMENT

The aim of this research was to evaluate subjects' judgment of sentence difficulty. Two groups of subjects were used, different patterns of sentences were arranged in pairs, each representing two of the four prescribed types, and subjects were required to select the harder sentence of each pair. Responses were evaluated along two dimensions: differences in sentence types and differences in subject group composition.

Samples

In order to avoid the problems of comparing two diverse samples, which were present in the De Pierro study, here research was restricted to enrollees in basic English courses at one school, Mercer County Community College (MCCC) in West Windsor, New Jersey. At matriculation, all students had been required to take the Comparative Guidance and Placement Program (CGP), a battery of tests assessing background, abilities and interests, developed for the College Entrance Examination Board by the Educational Testing Service. According to C. Robert Pace and H. Bradley Sagen in Buros, though this instrument has its philosophic weaknesses and lacks adequate validity and reliability data, it is well designed and normed, and is as "likely to perform about as effectively as other programs now available." (p. 1030)

The Guidance Department of MCCC uses this test to determine appropriate placement of students in introductory English classes. Students who score 43 or below (on a scale of 20-80) on the reading section are placed in Basic Composition (ES100); other students are advised to take Language and Literature (EG101). Subjects for this research were three classes of EG101 students, to be identified hereafter as "the good readers," and three classes of ES100 students, "the poor readers," for this experiment.

It should be noted that the original design of this research called for a single administration of the sentence judgment questionnaire, preferably at the same time as the administration of the CGP. However, this did not prove to be feasible. The next plan called for running the study at the beginning of the first summer session in the regularly scheduled classes. Subjects were obtained and the experiment conducted at this time, but there were not enough enrollees in the appropriate courses to have large enough samples. Therefore, the research had to be conducted during the second session as well. It is believed that no student filled out more than one questionnaire though.

In all, 105 subjects were finally obtained. Of these, 52 were categorized as poor students and 53 as good students. One potential subject's responses have not been included here because his knowledge of English was so poor; he could not understand what was expected of him. Another enrollee whose judgments were included turned out to be an eighth grade student who had been accepted as a summer school student and had demonstrated reading ability in line with the average of the class.

One final note on the poor readers: students enrolled in Basic Composition had been given the Test for Adult Basic Education, admittedly an instrument designed for children, not adults, which includes no data on reliability or norms (Buros, 60-63), and in the past, such students typically demonstrated reading levels between fourth and eighth grade on this test. Test results from one ES100 class were made available to the researcher and scores were found to range from 4.2, for a recent immigrant from Korea, to 11.0. (See Table 1.)

Sentence Production and Selection

One half of the sentence pairs used in this study (16) were those developed for the undergraduate sample of the De Pierro study. An additional sixteen pairs were developed in order to improve the statistical significance of the obtained data. Criteria for the new sentences were similar to those of the earlier study: 14 to 22 words in length with from 24 to 38 syllables, i.e., not too easy or too difficult for the subjects to react to the task. For the benefit of the poor readers, some attempt was made to simplify the vocabulary of some sentence selections.

Sources for the sentences in the De Pierro study are listed in Chapter II. The new sentences were derived from two texts recommended for eleventh and twelfth grades: English Literature and New Worlds of Literature. The words used in the two versions of each sentence pair were exactly the same in number and meaning as was the phrase structure, except for the intentionally created differences. Every attempt was made to minimize affective or semantic differences which often arise with position shifting.

TABLE 1

GRADE SCORES OF ONE ES100 CLASS ON TEST OF ADULT BASIC EDUCATION - JULY 1976

BASIC COMPOSITION - ES100 SECTION 9B24, MCCC
(N = 24)

11.0
10.1
10.0
9.5
9.1
8.5
8.4
8.1
8.0
8.0
7.9
7.8
7.7
7.6
7.6
7.5
7.2
6.7
6.3
6.2
6.2
6.2
5.7
4.2

Mean = 7.73

Standard deviation = 1.55

Sentences were manipulated to provide the four types of sentence distance suggested as crucial to comprehension in the Kernel Distance Theory and in the previously stated hypotheses:

D1: Distance at the beginning of the sentence before the kernel

D2: Distance between the subject and the verb

D3: Distance between the verb and the object

D4: Distance after the kernel, i.e., at the end of the sentence

In order to derive all four types from sentences found in published materials, substantial changes in original sentence structure, length, and lexicon were necessary. Only two variations were derived for each sentence to produce a pair for judgment. The final 32 pairs represent an equal number of comparisons (8) of the different distance positions under study, i.e., D1-D4, D2-D4, D2-D1, and D2-D3. Since there were 105 subjects, a total of 840 judgments were made for each type of comparison, 416 by the poor readers and 424 by the good readers.

Presentation

Sentence pairs were alternated so that one sentence variation, e.g., D1, was presented first half of the time and second, the other half of the time. All sentence pairs were then scrambled and typed, 4 pairs to a page, on 8 1/2" x 11" paper for duplication. Two forms of the questionnaire were developed and copies distributed alternately. It was hoped that this procedure would help to cancel possible effects of order of sentence presentation along with discouraging subject tendencies to "borrow" the judgments of adjacent students. An answer sheet was developed to match up to the arrangement of the sentence pairs.

Students were required to enter their test form, A or B, and the English course in which they were enrolled, along with their judgment of the sentence pairs.

Before questionnaires were distributed in each class, the researcher gave a brief oral explanation of the purposes and procedures of the study. First, readability was explained in very general terms; and it was emphasized that by their accurate responses to the questionnaire, their judgments would be of great use to textbook writers so students would ultimately have an easier time. Once the forms were distributed, instructions on the cover were read and time allowed for questions before anyone was allowed to begin. At this time, it was particularly emphasized that students should choose which sentence of each pair they found the harder. They were told to make a decision for every sentence pair and if they could not make up their minds, just to give a "gut reaction." Students were to check off their judgment of the harder sentence in the appropriate blank on the answer sheet. They were told that there would be no time limit and that there were no "correct" responses.

Student Responses

Various types of scaling techniques had been investigated, but because of the number of subjects involved, and the complex problems of interpretation that might result from sophisticated measures, it was decided to require a simpler judgment, a check-off in the appropriate blank of the harder sentence. When responses were tallied, it was found that, in fact, every student had made a judgment of every sentence pair. The questionnaire was completed in fifteen minutes on the average, although in a few cases, subjects finished in under ten minutes, and one subject took over half an hour.

Statistical Evaluation

After administration of the questionnaire, responses to Form B were converted to their Form A equivalent and sentence order for each comparison made the same. For example, now all D2-D4 and D4-D2 comparison responses were converted to the D2-D4 format so that tallying of sentence types would be accurate. Now all D2 sentences would be considered Sentence A, and the D4 sentences would be Sentence B. Statistical evaluation was then possible. It should be noted that responses of the good and poor readers were initially evaluated separately but later compared and combined.

Disparity of judgment of difficulty of different sentence types was analyzed using the X^2 test of goodness of fit:

$$X^2 = \frac{\Sigma(O - E)^2}{E}$$

where O = observed score, i.e., number who judged Sentence A to be harder and who judged Sentence B harder

E = expected score, i.e., half the number of possible judgments

In order to evaluate whether the differences in judgment between the good reader and poor reader samples were statistically significant, the X^2 test of homogeneity of two independent random samples was employed. The test statistic is the same as above, although this is a two sample test. Alpha was set at .05, giving a 95% level of confidence for all compilations.

In order to use these X^2 tests, it was necessary to temporarily convert all hypotheses to the null. Otherwise it would be impossible to determine an expected score. Since all calculations involved only one degree of freedom, the null hypothesis was rejected (at the .05 level) if X^2 equalled 3.84 or greater, i.e., if the difference in proportion of the two samples was statistically significant. Directionality was then determined by inspection of the raw data and then compared to the original directional hypothesis.

The X^2 tests utilized make several assumptions:

1. It is assumed that the sample is drawn randomly from the population. This is probably the hardest factor to achieve in experiments in the behavioral sciences. It would have been realized best for this study if an unlimited number of students had been available and smaller samples for the research drawn at random. However, as explained earlier, this was not possible. Therefore, results will only be applicable to similar samples to those utilized here.
2. Observations are assumed to be independent. This was insured by the experimental design. If a subject indicated that one sentence of a pair was harder than another, only one score was entered in one of the two possible cells, i.e., if he said Sentence A was harder, for that item A received a score of one; Sentence B was scored as a 0.

3. It is assumed that in repeated experiments observed, frequencies will be normally distributed about expected frequencies. In order to enhance this possibility, the number of sentence pairs was doubled and sample size expanded to over fifty for each group.

Actual calculations were done on an HP-55 pocket calculator utilizing the appropriate programs provided with the Hewlett Packard HP-55 Statistics Programs (1974) after their accuracy had been verified.

CHAPTER IV

RESULTS

In this chapter, the results of this investigation are presented. Each hypothesis is considered individually in order of presentation in Chapter I. Results are then briefly summarized, more detailed discussion being reserved for the final chapter.

Hypothesis 1 (D1 > D4)

Hypothesis One predicted that sentences with distance at the beginning would be judged harder than distance at the end. As may be seen in Table 2, among the poor readers, the good readers and when both groups were combined, no statistically significant differences were found. Judgments of individual sentence pairs in a particular category, such as Item 32 on the questionnaire, in which the D4 case was judged much harder than the D1 sentence, were off-set by other sentence pairs in which D1 was considered harder than D4. The comparison of the two reader groups shows that the poor readers had a tendency to find the D1 sentence of Item 32 harder than the D4 sentence, but this was off-set when considering total reader judgments by the responses of the good readers.

In other words, in the case of this hypothesis, any great disparity that manifested itself on an individual item was off-set by judgments on other items in this comparison category; so much so, in fact, that by the time all responses of good and poor readers

TABLE 2

TALLIES OF STUDENT JUDGMENTS OF MORE DIFFICULT SENTENCES $H_o 1$: D1 $>$ D4
(N = 105 students)

Item Number		Poor Reader Judgments		Good Reader Judgments		Total Reader Judgments	
Form A	Form B	D1	D4	D1	D4	D1	D4
1	32	26	26	29	24	55	50
7	26	29	23	21	32	50	55
13	20	27	25	31	22	58	47
17	16	24	28	24	29	48	57
20	13	21	31	32	21	53	52
26	7	27	25	33	20	60	45
30	3	27	25	19	34	46	59
32	1	32	20	16	37	48	57
Totals		213	203	205	219	418	422
Statistical Results		$X^2 = .2403$		$X^2 = .4623$		$X^2 = .0191$	
Decisions		Do not reject null H_o		Do not reject null H_o		Do not reject H_o	

Decision Rule: When $\alpha = .05$ reject H_o if $X^2 \geq 3.84$

were totalled, out of 840 judgments, 418 judged distance at the beginning harder than sentences with distance at the end and 422 judged the reverse. This is practically an even split. However, as will be seen, judgments of other types of sentence pairs point to greater difficulty for one type as compared with another.

Hypothesis 2 (D2 > D4)

Hypothesis Two predicted that sentences with distance between subject and verb would be judged harder than those with distance at the end of a sentence. The raw data and X^2 results may be found for this sentence comparison on Table 3. In this case, greater disparity resulted. More poor readers and good readers judged the D2 type of sentence harder than the D4 type. Within each reading group, this difference was not quite high enough to be significant; but when the reading group scores were totalled, a statistically significant difference in judgments was found, thereby confirming the original hypothesis.

Hypothesis 3 (D2 > D1)

The third hypothesis predicted that sentences with distance between subject and verb would be judged more difficult than those with distance at the beginning of a sentence. With this type of sentence comparison, a clear indication of difference in difficulty was found, as may be seen in Table 4. Poor readers and especially good readers judged the D2 sentences harder in a statistically significant number of cases. When scores of both groups were totalled, of 840 judgments, close to two-thirds found the D2 sentences more difficult. The results confirm Hypothesis Three.

TABLE 3

TALLIES OF STUDENT JUDGMENTS OF MORE DIFFICULT SENTENCES H_o2: $D2 \geq D4$

(N = 105 students)

Item Number		Poor Reader Judgments		Good Reader Judgments		Total Reader Judgments	
Form A	Form B	D2	D4	D2	D4	D2	D4
3	30	25	27	17	36	42	63
4	29	31	21	26	27	57	48
8	25	28	24	31	22	49	46
10	23	31	21	28	25	59	46
16	17	25	27	34	19	59	46
21	12	33	19	37	16	70	35
24	9	34	18	42	11	76	29
29	4	18	34	10	43	28	77
Totals		225	191	225	199	450	390
Statistical Results		$X^2 = 2.77$		$X^2 = 1.594$		$X^2 = 4.286$	
Decisions		Do not reject null H_o		Do not reject null H_o		Reject null H_o	

Decision Rule: When $\alpha = .05$ reject H_o if $X^2 \geq 3.84$

TABLE 4

TALLIES OF STUDENT JUDGMENTS OF MORE DIFFICULT SENTENCES H_o3: D2 > D1
(N = 105 students)

Item Number		Poor Reader Judgments		Good Reader Judgments		Total Reader Judgments	
Form A	Form B	D1	D2	D1	D2	D1	D2
2	31	21	31	18	35	39	66
5	28	15	37	12	41	27	78
6	27	31	21	33	20	64	41
12	21	12	40	6	47	18	87
15	18	10	42	13	40	23	82
18	15	27	25	34	19	61	44
23	10	20	32	9	44	29	76
28	5	21	31	10	43	31	74
Totals		157	259	135	289	292	548
Statistical Results		$x^2 = 25.009$		$x^2 = 55.934$		$x^2 = 78.019$	
Decisions		Reject null H_o		Reject null H_o		Reject null H_o	

Decision Rule: When $\alpha = .05$ reject H_o if $x^2 \geq 3.84$

Hypothesis 4 (D2 > D3)

Hypothesis Four predicted that sentences with distance between subject and verb would be judged more difficult than those with distance between verb and object. The raw data and statistical results may be found on Table 5. The pattern of results for this hypothesis is similar, although somewhat more exaggerated than that for Hypothesis One. Poor readers judged the D2 sentences harder than the D3 sentences, although not statistically significantly so; the good readers found the D3 sentences more difficult than the D2 sentences, but, again, not significantly. The total scores of the combined groups, therefore, balance each other out, resulting in an overall judgment of virtually no difference in difficulty between D2 and D3 type sentences.

Hypothesis 5: Reading Group Comparisons

Hypothesis Five predicted that subjects with higher reading scores would judge sentence difficulty differently from subjects with low reading scores. Raw data totals for each reading group obtained in the earlier analyses is repeated in Table 6, along with the obtained scores from the appropriate X^2 tests of homogeneity. It is apparent that no statistically significant difference was found between the two reading groups for the first three hypotheses, contradicting Hypothesis Five. However, in the case of the Fourth hypothesis, a significant difference between reading group responses was found: the good readers judged D3 type sentences harder than the D2 type, and the poor readers reported the reverse. In only this case was Hypothesis Five confirmed.

TABLE 5

TALLIES OF STUDENT JUDGMENTS OF MORE DIFFICULT SENTENCES H_o4: D2 > D3
(N = 105 students)

Item Number		Poor Reader Judgments		Good Reader Judgments		Total Reader Judgments	
Form A	Form B	D2	D3	D2	D3	D2	D3
9	24	25	27	15	38	40	65
11	22	26	26	17	36	43	62
14	19	34	18	35	18	69	36
19	14	24	28	26	27	50	55
22	11	18	34	7	46	25	80
25	8	38	14	45	8	83	22
27	6	31	21	34	19	65	40
31	2	26	26	15	38	41	64
Totals		222	194	194	230	416	424
Statistical Results		$X^2 = 1.885$		$X^2 = 3.057$		$X^2 = .076$	
Decisions		Do not reject null H_o		Do not reject null H_o		Do not reject null H_o	

Decision Rule: When $\alpha = .05$ reject H_o if $X^2 \geq 3.84$

TABLE 6

COMPARISONS OF GOOD AND POOR READER JUDGMENTS OF MORE DIFFICULT SENTENCES

HYPOTHESES ONE, TWO, THREE AND FOUR
(N = 105 students)

Sample	Hypothesis One		Hypothesis Two		Hypothesis Three		Hypothesis Four	
	D1	D4	D2	D4	D1	D2	D2	D3
Good Readers	205	219	225	199	135	289	194	230
Poor Readers	213	203	225	191	157	259	222	194
Statistical Results	$X^2 = .57$		$X^2 = .05$		$X^2 = 2.97$		$X^2 = 4.57$	
Decisions	Do not reject null H_o		Do not reject null H_o		Do not reject null H_o		Reject null H_o	

Decision Rule: When $\alpha = .05$ reject H_o if $X^2 \geq 3.84$

Summary of the Results

Although subject judgments on individual items may vary widely within sentence categories, when these were totalled, certain trends emerged. In two of the four types of sentence comparisons (Hypotheses Two and Three), some significant difference in difficulty was reported and in the direction predicted by the hypotheses. For three out of the four hypotheses, good and poor readers did not judge sentences significantly differently. However, in the one significant case (Hypothesis Four) and also by inspection of the raw data of Hypothesis One, one notes some tendency on the part of the good readers to judge sentences in the opposite direction from that predicted.

CHAPTER V

DISCUSSION AND CONCLUSIONS

Comparison With Earlier Research

Since the only other investigation to validate the Kernel Distance Theory has been the De Pierro study, this section will concentrate on a comparison with this earlier research. Possible confirmations or negations of other studies and/or theories will only be considered in the context of this central discussion.

Table 7 summarizes the findings of De Pierro's study of sentence structure variables on three measures of reading ease, along with the findings of the current research using student judgment of sentences. The latter investigation confirms some of the earlier results and contradicts others; neither appears to provide clear-cut validation of all of the Kernel Distance Theory.

Hypothesis One predicted that sentences with distance at the beginning would be judged more difficult than those with distance at the end. It was assumed, as predicted by the Yngve model (1960), that inserting non-essential material to be stored before the main kernel of the sentence was presented would strain the memory system, resulting in a harder sentence. This was not found to be the case in either the De Pierro study or the current research.

TABLE 7

COMPARISON OF RESULTS OF TWO STUDIES BASED ON THE KERNEL DISTANCE THEORY

	READING EASE STUDY (DE PIERRO)						STUDENT JUDGMENT STUDY (WEBER)		
	Elementary Subjects			College Subjects			Reading Ability		
	Recall	Silent Reading Time	Response Delay	Recall	Silent Reading Time	Response Delay	Poor	Good	Total Subjects
H_o1 $D1 > D4$	$0_{(-)}$	0	0	-	-	0	0	$0_{(-)}$	0
H_o2 $D2 > D4$	+	+	0	0	+	0	0	0	+
H_o3 $D2 > D1$	+	0	+	+	$0_{(-)}$	0	+	+	+
H_o4 $D2 > D3$	0	0	0	0	0	0	0	$0_{(-)}$	$0_{(-)}$

0 = Non-significant Difference
+ = Significant Difference in Predicted Direction
- = Significant Difference in Non-predicted Direction
( ) = Indicates Tendencies, though Non-significant

De Pierro found no significant difference between such sentences with the elementary school sample, although a tendency in the direction opposite to that expected, finding sentences with distance at the end to be harder, was noted in the recall measures. This contradiction became more pronounced with De Pierro's college group both on measures of recall and silent reading time. The new investigation noted a tendency towards this contradiction among the good readers, but no significant difference was found between these two types of sentences.

De Pierro suggests that readers who find no significant difference between such sentences may not perceive syntactic difference between them, since the main constituents of the sentence, subject, verb and object, are not split. Readers who find the sentences with distance at the end harder than those with distance at the beginning, may skim over a non-essential chunk at the end of a sentence, which would probably not be done if it were at the beginning. This would certainly account for De Pierro's findings of poorer recall measures on such sentences, but would probably not influence students' judgment of a sentence if they were not required to do anything else with it.

It should be noted here that though these findings are contradictory to Yngve's model, they do not necessarily support any other model. For example, transformational grammar may not be applicable here since the sentence manipulations of this hypothesis did not involve major transformational shifts.

Hypothesis Two predicted great difficulty for sentences with a subject-verb split than those with distance at the end. De Pierro found this to be true, more so with his younger sample. The present study also found overall confirmation of this hypothesis although it was not apparent in the independent analyses of the separate reading groups.

Hypothesis Three was similar to Hypothesis Two, comparing sentences with subject-verb splits with sentences with no constituent splits but distance at the beginning. The De Pierro findings were generally supportive, especially with the younger group and the current research found this to be true to a very high degree among all groups, in strong support of the Kernel Distance Theory. It seems plausible that this indication should have been so much stronger for the good readers than that found in their judgment of the Hypothesis Two sentences since Hypothesis One had previously shown that this group judged sentences with distance at the beginning easier than those with distance at the end. Poorer readers probably find any major constituent split, especially the subject-verb split, harder to deal with than a chunk anywhere outside the kernel. It is not clear why De Pierro's measures were not more positive.

The results of Hypotheses Two and Three are in accordance with those of earlier researchers who studied subject-verb splits. Coleman (1964) presented a college sample with sentence pairs on a memory drum and measured numbers of trials necessary for perfect memorization. He found that "embedded" sentences, as he called those with distance between subject and verb, harder to learn than non-embedded sentences. Hamilton and Deese (1971) using similar sentence comparisons in a study of listening comprehension, found "center-embedded" sentences (their terminology for subject-verb splits) to be more difficult to comprehend. These findings are also in accord with the syntactic measures of Marcus (1971) and Botel and Granowsky (1972).

Coleman also notes that sentences of this type are beyond the capacity of most machine programs which "create language." This is true since most of these programs have been designed according to the rules of finite-state grammar.

Hypothesis Four predicted that subject-verb splits would be judged more difficult than verb-object splits. All De Pierro's measures found such splits to be equally difficult for both his groups. The subject judgment study found some tendency among the good readers to judge the verb-object splits harder, although at a non-significant level; the poor readers performed more closely to the subjects of the earlier study. It is unclear why this occurred. Perhaps the good readers, too, perceived these sentences as being equally difficult, but since such a judgment could not be indicated by the experiment, these subjects may have arbitrarily chosen the verb-object splits over the subject-verb splits without any thought.

Little earlier research comparing such sentences was found in the literature. Yngve's model had predicted that the verb-object splits would be found easier than the subject verb splits. Both studies found no support for this prediction.

In summation, both investigations give some support to Fry's Kernel Distance Theory but other parts are more suspect. Sentences with subject-verb splits appear to be harder than sentences with non-essential distance outside the kernel. Subjects when stored before they can be meaningfully linked with the verb may begin to fade, creating a weaker memory connection and consequently, a more difficult sentence.

Different kinds of kernel splits, whether subject-verb or verb-object appear to be equally difficult, perhaps because memory connections are as important between verb and object as between subject and verb. As for distance at the beginning compared to distance at the end of a sentence, in spite of theoretical predictions, findings thus far should be considered ambiguous at best.

General Applications

The wider aim of this research was to seek possible syntactic indicators of reading ease or complexity for the benefit of writers, revision editors, and future measures of readability. On the basis of this study, it seems reasonable to advise that language, whether written or oral, which includes constituent splits, whether between subject and verb or between verb and object, should be avoided if reading ease is of concern. However, it should not be assumed that this will automatically insure greater oral comprehension or readability; in fact, it may have very little impact.

As has been indicated earlier, readability is a highly complex phenomenon influenced by both objective and subjective factors, some of which may be controlled by writers, but many of which cannot. The sentence pairs used in this study were presented out of any context, which might be expected to influence interest and provide decoding clues, syntactic and semantic. Affective factors, such as subjects' backgrounds and interests, which are brought to the reading situation, have also been shown to be highly influential.

As for readability formulas, this researcher must agree at this time with De Pierro. Given the number of factors which influence readability, it is probably inappropriate for a formula to be sensitive to all of them. Those that do try to take many syntactic factors into consideration are cumbersome to use and, therefore, are not practical for use by classroom teachers. For the present, perhaps it is best to rely on indicators of sentence complexity, such as sentence length, until something better becomes available. Hopefully, researchers will continue to probe syntactic variations and their relation to other variables of language to improve our understanding and measures of readability.

Suggestions for Further Research

Continued investigation of the influence of syntax on language appears to be warranted, using contrasting patterns in isolation, as in this study, as well as in context. The problem of choosing appropriate dependent variables for measuring ease of comprehension remains. In this research, subject judgments were utilized. Students appeared to be quite pleased that they were being consulted for their opinion and consequently many apparently took the questionnaire seriously. Some, however, were bothered that they could not indicate a "no difference" judgment. Future studies of this nature might consider using scaled responses rather than the simple "which is harder" technique used here. How arbitrarily answers are made and how seriously the task is completed can never be fully controlled and studies utilizing this procedure should always be viewed accordingly. Using large numbers of subjects and questionnaire items, hopefully, counter-balances these problems. An interesting, motivating introduction explaining the importance of the study to the subjects is also essential.

Future research should also take linguistic background of subjects into consideration. Students used in the investigation, described here, were discovered to have varying linguistic backgrounds; some were found to be non-native speakers of English and others spoke a non-standard dialect. Judgments of such speakers might be analyzed separately and in relation to the rules of the appropriate dialects for a clearer picture of why some types of sentence constructions might be found harder for some subjects but not others.

Future investigations should also consider narrowing down the sample along other lines. The current research classified reading ability in one grade using the categories established by the participating college. However, wide variability in ability was still found. Ideally, samples with narrower ranges of abilities should be chosen.

Future studies must also continue to provide tight control over experimental syntactic patterns. Earlier research has noted problems of maintaining semantic equivalency and length control. In the current investigation and the De Pierro study length of sentences and words used for each sentence of a pair were identical. Phrase structure was also closely matched. It is suggested that such practices be continued where applicable.

Additional investigations should be done on kernel splits, especially on verb-object splits which were not thoroughly compared to other patterns here. Further analysis of distance at the beginning and end of sentences is also warranted. Overall complexity could also be varied to assess which types of more complex sentences might be easier than simpler patterns, as suggested by the Pearson research (1975).

On a broader level, there is much need for research on the interaction of syntactics and semantics in isolation as well as in the context of larger passages than sentences. Ultimately, the narrower studies suggested here should also be applied to a more developmental approach within and between dialect groups.

REFERENCES

Bormuth, J. New data on readability. Address before a joint meeting of the International Reading Association and the American Educational Research Association. Seattle: 1967. ED016586.

Bormuth, J. Readability: a new approach. Reading Research Quarterly, 1966, 1, 79-132.

Botel, M. and Granowsky, A. Formula for measuring syntactic complexity. Elementary English, 1972, 49, 513-516.

Bransford, J. D. and Franks, J. J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.

Buros, O. K., edited by. The Seventh Mental Measurements Yearbook. Highland Park, N. J.: The Gryphon Press, 1972.

Carroll, J. B. Measurement properties of subjective magnitude estimates of word frequency. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 722-729.

Carver, R. P. Understanding information processing and learning from prose materials. Journal of Educational Psychology, 1973, 64, 76-84.

Chall, J. S. Readability. Columbus, Ohio: Ohio State University, 1958.

Chomsky, N. Aspects of the Theory of Syntax. Cambridge, Mass.: M.I.T. Press, 1965.

Chomsky, N. Syntactic Structures. The Hague: Mouton, 1957.

Cline, T. A. Readability of community college textbooks and the reading ability of the students who use them. Journal of Reading Behavior, 1972-1973, 5, 110-118.

Coleman, E. B. Learning of prose written in four grammatical transformations. Journal of Applied Psychology, 1965, 49, 332-341.

Coleman, E. B. The comprehensibility of several grammatical transformations. Journal of Applied Psychology, 1964, 48, 186-190

Daiches, D., Jewett, A., Havighurst, W. and Searles, J. English Literature. Boston: Houghton Mifflin Company, 1965.

Dale, E. and Tyler, R. W. A study of the factors influencing the difficulty of reading materials for adults of limited reading ability. Library Quarterly, 1934, 4, 384-412.

De Pierro, J. P. Some effects of sentence structure variables on reading ease. Unpublished doctoral dissertation. New Brunswick, N. J.: Rutgers University, 1976.

Estes, T. H. Scale to measure attitudes toward reading. Journal of Reading, 1971, 15, 135-138.

Flesch, R. A new readability yardstick. Journal of Applied Psychology, 1948, 32, 221-233.

Flesch, R. How to Test Readability. New York: Harper & Brothers, 1951.

Fodor, J. A., Garrett, M. and Bever, T. G. Some syntactic determinants of sentential complexity, II: verb structure. Perception and Psychophysics, 1968, 3, 453-461.

Fries, C. C. Linguistics and Reading. New York: Holt, Rinehart and Winston, 1963.

Fry, E. A kernel distance theory for readability. In G. H. McNinch and W. D. Miller (Eds.), Reading: Convention and Inquiry. Twenty-Fourth Yearbook of the National Reading Conference. Clemson, South Carolina: The National Reading Conference, Inc., 1975.

Gazer, S. M. Is sentence length a valid measure of difficulty in readability formulas? Reading Teacher, 1974, 27, 464-468.

Goodman, K. S. Reading: a psycholinguistic guessing game. Reprinted in H. Singer and R. B. Ruddell (Eds.), Theoretical Models of Reading. Newark, Delaware: International Reading Association, 1970, 259-272.

Gough, P. B. Grammatical transformations and speed of understanding. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 107-111.

Gray, W. S. and Leary, B. E. What Makes a Book Readable? Chicago: University of Chicago Press, 1935.

Greenwald, A. G. On defining attitude and attitude theory. In A. G. Greenwald, T. C. Brock and T. M. Ostrom, Psychological Foundations of Attitudes. New York: Academic Press, 1968, 361-390.

Guthrie, J. T. Reading comprehension and syntactic responses in good and poor readers. Journal of Educational Psychology, 1973, 65, 294-299.

Hagstrom, J. M. A comparison of the reading abilities of a junior college occupational education population and the readability levels of their texts. A vocational education project under Title I, Part C, Sec. 131 (b) of PL60-576 and CRCC Small Grant Projects. Columbia Junior College, Calif., 1973. ED088028.

Halliburton, W. J. and Pelkonen, M. E. New Worlds of Literature, New York: Harcourt, Brace and World, Inc., 1966.

Hamilton, H. W. and Deese, J. Comprehensibility and subject-verb relations in complex sentences. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 163-170.

Harrell, L. E., Jr. A comparison of the development of oral and written language in school-age children. Monographs of the Society for Research in Child Development, 1957, 22, No. 3.

Harris, A. J. Some new developments in readability. Paper presented at the annual International Reading Association World Congress on Reading. Vienna, Austria, August 12-14, 1974. ED094344.

Herriot, P. Comprehension of sentences as a function of grammatical depth and order. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 938-941.

Hittleman, D. R. Seeking a psycholinguistic definition of readability. Reading Teacher, 1972-73, 26, 783-789.

Hoffner, Dorothea R. A Psycholinguistic analysis of oral reading miscues by junior college students. Unpublished Master's Dissertation. New Brunswick, N. J.: Rutgers University, 1974.

Howe, E. S. Transformation, associative uncertainty, and free recall of sentences. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 425-430.

Hulse, S. H., Deese, J. and Egeth, H. The Psychology of Learning. New York: McGraw-Hill Book Company, 1975, 281-324.

Hunt, K. W. Recent measures in syntactic development. In M. Lester (Ed.), Readings in Applied Transformational Grammar. New York: Holt, Rinehart and Winston, Inc., 1970, 187-200.

Jorgenson, G. W. Analysis of teacher judgments of reading level. American Educational Research Journal, 1975, 12, 67-75.

Klare, G. R. Assessing readability. Reading Research Quarterly, 1974-1975, 10, 62-102

Klare, G. R. The Measurement of Readability. Ames, Iowa: Iowa State University Press, 1963.

Klare, G. R. and Buck, B. Know Your Reader. New York: Hermitage House, 1954.

Kolers, P. A. Three stages of reading. In Smith, F., Psycholinguistics and Reading. New York: Holt, Rinehart and Winston, Inc.,1973.

Lewerenz, A. S. Measurement of the difficulty of reading materials. Los Angeles Educational Research Bulletin, 1929, 8, 11-16.

Lippman, M. Z. The influence of grammatical transform in a syllogistic reasoning task. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 424-430.

Lively, B. A. and Pressey, S. L. A method for measuring the "vocabulary burden" of textbooks. Educational Administration and Supervision, 1923, 9, 389-398.

Loban, W. D. The Language of Elementary School Children. (Research Report No. 1) Champaign, Illinois: National Council of Teachers of English, 1969.

Lorge, I. Predicting reading difficulty of selections for children. Elementary English Review, 1939, 16, 229-233.

MacGinitie, W. H. and Tretiak, R. Sentence depth measures as predictors of reading difficulty. Reading Research Quarterly, 1971, 6, 364-377.

Marcus, A. D. The development of a diagnostic test of syntactic meaning clues in reading. In R. E. Leibert (Ed.), Diagnostic Viewpoints in Reading. Newark, Delaware: International Reading Association, 1971, 48-63.

Mehler, J. Some effects of grammatical transformations on the recall of English sentences. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 340-351.

Mehler, J. and Carey, P. The interaction of veracity and syntax in the processing of sentences. Perception and Psychophysics, 1968, 3, 109-111.

Miller, G. A. Some psychological studies of grammar. American Psychologist, 1962, 17, 748-762.

Minium, E. W. Statistical Reasoning in Psychology and Education. New York: John Wiley and Sons, Inc., 1970.

Moir, H. Linguistic factors related to style and meaning in written language. Elementary English, 1970, 47, 215-222.

Ohaver, R. A comparison study of semantic and syntactic cueing by low reading performance college freshmen. Tampa, Fla.: Address delivered before a meeting of the National Reading Conference, 1971. ED059836.

Ojemann, R. The readability of parents and factors associated with the reading difficulty of parent education materials. In G. D. Stoddard (Ed.) Researches in Parent Education II. Iowa City, Iowa: University of Iowa, 1934.

Pearson, P. D. The effects of grammatical complexity on children's comprehension, recall, and conception of certain semantic relations. Reading Research Quarterly, 1974-1975, 10, 155-192.

Perfetti, C. A. Lexical density and phrase structure depth as variables in sentence retention. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 719-724.

Rohrman, N. L. The role of syntactic structure in the recall of English nominalizations. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 348-353.

Ruddell, R. B. An investigation of the effect of the similarity of oral and written language structure on reading comprehension. Unpublished doctoral dissertation. Indiana University, 1963.

Ruddell, R. B. Psycholinguistic implications for a systems of communication model. In H. Singer and R. B. Ruddell (Eds.) Theoretical Models and Processes of Reading. Newark, Delaware: International Reading Association, 1970, 239-258.

Shapiro, B. J. The subjective estimation of relative word frequency. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 248-251.

Slobin, D. I. Grammatical transformations and sentence comprehension in childhood and adulthood. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 219-227.

Smith, F. Psycholinguistics and reading. In F. Smith (Ed.) Psycholinguistics and Reading. New York: Holt, Rinehart and Winston, Inc., 1973.

Spring, D. S. How much do community college students learn from their textbooks? Journal of Reading, 1975, 19, 131-136.

Stevens, S. S. The direct estimation of sensory magnitudes - loudness. American Journal of Psychology, 1956, 69, 1-25.

Strang, R. Estimating the difficulty of high-school and college reading material. (Abstract) Practical Values of Educational Research. Washington: American Educational Research Association, 1938.

Strickland, R. The language of elementary school children. Bulletin of the School of Education, 1962, 38.

Tatham, S. M. Reading comprehension of materials written with select oral language patterns: a study at grades two and four. Unpublished doctoral dissertation. Madison: University of Wisconsin, 1969.

Taylor, W. L. Cloze procedure: a new tool for measuring readability. Journalism Quarterly, 1953, 30, 415-433.

Thorndike, E. L. Reading as reasoning: a study of mistakes in paragraph reading. Journal of Educational Psychology. 1917, 8, 323-332.

Washburne, C. and Vogel, M. An objective method of determining grade placement of children's reading material. Elementary School Journal, 1928, 373-381.

Wood, R. W. and Schlief, M. Comparison of procedures to determine readability level on non-text materials. Reading Improvement, 1974, 11, 57-64.

Yngve, V. H. A model and an hypothesis for language structure. Proceedings of the American Philosophical Society, 1960, 104, 444-466.

APPENDIX A

APPENDIX A

WEBER SENTENCES

H_o1: D1 > D4

SENTENCE VERSIONS D1 AND D4

	D1	D4
A)	After a delay that seemed to last forever, the train finally moved out of the station slowly.	The train finally moved out of the station slowly after a delay that seemed to last forever.
B)	Although he first studied art, in his early twenties Chesterton became a practicing journalist.	In his early twenties Chesterton became a practicing journalist although he first studied art.
C)	Made possible by changes in the traditional philosophy of life, comfort is now one of the causes of its own further spread.	Comfort is now one of the causes of its own further spread, made possible by changes in the traditional philosophy of life.
D)	Shortly after the orchestra began playing a marching song, the people gathered around the candidate.	The people gathered around the candidate shortly after the orchestra began playing a marching song.

WEBER SENTENCES

H_o2: D2 > D4

SENTENCE VERSIONS D2 AND D4

	D2	D4
A)	They, having come on foot over several hundred miles, were a very tired looking people.	They were a very tired looking people, having come on foot over several hundred miles.
B)	We waited breathlessly as the queen, followed by three attendants, walked gracefully out on the court.	We waited breathlessly as the queen walked gracefully out on the court, followed by three attendants.
C)	A Scottish accent, for example, has never been considered a social handicap in this country.	A Scottish accent, has never been considered a social handicap in this country, for example.
D)	The stranger in the corner, waving his arms in time with the music, splashed his beer all over the floor.	The stranger in the corner splashed his beer all over the floor waving his arms in time with the music.

WEBER SENTENCES

$H_o 3$: $D2 \lambda D1$

SENTENCE VERSIONS D2 AND D1

	D2	D1
A)	The other residents of the neighborhood, the school board said, did not live in the same district as our block.	The school board said the other residents of the neighborhood did not live in the same district as our block.
B)	Tom Dooley, skilled in medicine and fluent in French, proved valuable working in the hills of Southeast Asia.	Skilled in medicine and fluent in French, Tom Dooley proved valuable working in the hills of Southeast Asia.
C)	The woman on welfare found, when the relative's will was read, that her monetary problems were over.	When the relative's will was read the woman on welfare found that her monetary problems were over.
D)	The reporters, as the troops raced over a rise in the land, finally caught sight of the enemy.	As the troops raced over a rise in the land the reports finally caught sight of the enemy.

WEBER SENTENCES

H_o4: D2 > D3

SENTENCE VERSIONS D2 AND D3

	D2	D3
A)	Someone with a trained ear, as Shaw's Professor Higgins has, can distinguish the region in which a speaker resides.	Someone with a trained ear can distinguish, as Shaw's Professor Higgins has, the region in which a speaker resides.
B)	I, when I went across the street to my apartment, discovered my best friend waiting impatiently.	I discovered, when I went across the street to my apartment, my best friend waiting impatiently.
C)	We, watching his sleepwalking, found we were most interested in where our little brother hid what he collected.	We found, watching his sleepwalking, we were most interested in where our little brother hid what he collected.
D)	All the people, before we had gone ten miles were crowded together on old-time open platforms.	All the people, were crowded together before we had gone ten miles on old-time open platforms.

APPENDIX B

APPENDIX B

DE PIERRO SENTENCES

H_o1: D1 > D4

SENTENCE VERSIONS D1 AND D4 FOR UNDERGRADUATE SAMPLE

	D1	D4
A)	By debating and agonizing over every problem, we show ourselves to be exceedingly timid and shy.	We show ourselves to be exceedingly timid and shy by debating and agonizing over every problem.
B)	When many individual observations and thoughts are pooled, human knowledge and vision are enhanced.	Human knowledge and vision are enhanced when many individual observations and thoughts are pooled.
C)	While motoring through the mottled countryside, the vacationer becomes impressed with the complicated farm machinery.	The vacationer becomes impressed with the complicated farm machinery while motoring through the mottled countryside.
D)	Unless they act by the insights they have, men will be no more the passive recipients of orders.	Men will be no more than the passive recipients of orders unless they act by the insights they have.

DE PIERRO SENTENCES

H_o2: D2 > D4

SENTENCE VERSIONS D2 AND D4 FOR UNDERGRADUATE SAMPLE

	D2	D4
A)	Some psychologists, by conducting carefully controlled studies and surveys, have produced valuable information.	Some psychologists have produced valuable information by conducting carefully controlled studies and surveys.
B)	No belief, if injustices and evils are to be eradicated, can be regarded as infallible.	No belief can be regarded as infallible if injustices and evils are to be eradicated.
C)	Suppressed groups, when reaching an historical stage of awareness, grab violently for independence.	Suppressed groups grab violently for independence when reaching an historical stage of awareness.
D)	Sickness and disease, by their constant yet unpredictable threat, introduce an element of tragic struggle in the world.	Sickness and disease introduce an element of tragic struggle in the world by their constant yet unpredictable threat.

DE PIERRO SENTENCES

H_o3: D2 > D1

SENTENCE VERSIONS D2 AND D1 FOR UNDERGRADUATE SAMPLE

	D2	D1
A)	The Renaissance artist, because of new emphasis on man and the universe, realistically depicted nature.	Because of new emphasis on man and the universe, the Renaissance artist realistically depicted nature.
B)	Baird, even during his daily business activities, had speculated with sending human images through space.	Even during his daily business activities, Baird had speculated with sending human images through space.
C)	The word speed, because of its derivation from the Anglo-Saxon word sped, does not denote recklessness.	Because of its derivation from the Anglo-Saxon word sped, the word speed does not denote recklessness.
D)	The troop decrease, as the evacuation rapidly commenced, was accompanied by a weakening of defenses.	As the evacuation rapidly commenced, the troop decrease was accompanied by a weakening of defenses.

DE PIERRO SENTENCES

H_o4: D2 > D3

SENTENCE VERSIONS D2 AND D3 FOR UNDERGRADUATE SAMPLE

	D2	D3
A)	Middle-Age painting, by serving as the Church's handmaiden, embellished the doctrines of Christianity.	Middle-Age painting embellished, by serving as the Church's handmaiden, the doctrines of Christianity.
B)	Sociologists, despite the challenges posed by higher education, have rarely conducted insightful studies of colleges.	Sociologists, have rarely conducted, despite the challenges posed by higher education insightful studies of colleges.
C)	The moralist, while reflecting upon the brevity of human life, has employed a variety of favorite symbols.	The moralist has employed, while reflecting upon the brevity of human life, a variety of favorite symbols.
D)	Humor and religion, by changing one's outlook and perspective, can precipitate anxieties into sane behavior.	Humor and religion can precipitate, by changing one's outlook and perspective, anxieties into sane behavior.

APPENDIX C

APPENDIX C

SAMPLE QUESTIONNAIRE - FORM A

THIS BOOKLET CONTAINS 32 PAIRS OF SENTENCES

BOTH SENTENCES OF EACH PAIR CONTAIN
THE SAME WORDS BUT THEIR ORDER
IS DIFFERENT

PLEASE INDICATE ON THE ANSWER SHEET
WHICH SENTENCE IN EACH PAIR
YOU THINK IS HARDER

THERE ARE NO RIGHT OR WRONG ANSWERS

THERE IS NO TIME LIMIT

THANK YOU FOR YOUR HELP!
Jane E. Weber
Rutgers University

Form A/1

1. A. Shortly after the orchestra began playing a marching song, the people gathered around the candidate.

   B. The people gathered around the candidate shortly after the orchestra began playing a marching song.

2. A. Even during his daily business activities, Baird had speculated with sending human images through space.

   B. Baird, even during his daily business activities, had speculated with sending human images through space.

3. A. The stranger in the corner splashed his beer all over the floor, waving his arms in time with the music.

   B. The stranger in the corner, waving his arms in time with the music, splashed his beer all over the floor.

4. A. No belief, if injustices and evils are to be eradicated, can be regarded as infallible.

   B. No belief can be regarded as infallible if injustices and evils are to be eradicated.

5.	A.	The Renaissance artist, because of new emphasis on man and the universe, realistically depicted nature.	B.	Because of new emphasis on man and the universe, the Renaissance artist realistically depicted nature.
6.	A.	Skilled in medicine and fluent in French, Tom Dooley proved valuable working in the hills of Southeast Asia.	B.	Tom Dooley, skilled in medicine and fluent in French, proved valuable in the hills of Southeast Asia.
7.	A.	Made possible by changes in the traditional philosophy of life, comfort is now one of the causes of its own further spread.	B.	Comfort is now one of the causes of its own further spread, made possible by changes in the traditional philosophy of life.
8.	A.	Suppressed groups grab violently for independence when reaching an historical stage of awareness.	B.	Suppressed groups, when reaching an historical stage of awareness, grab violently for independence.

9.	A.	Middle-Age painting, by serving as the Church's handmaiden, embellished the doctrines of Christianity.	B.	Middle-Age painting embellished, by serving as the Church's handmaiden, the doctrines of Christianity.
10.	A.	We waited breathlessly as the queen walked gracefully out on the court, followed by three attendants.	B.	We waited breathlessly as the queen, followed by three attendants, walked gracefully out on the court.
11.	A.	Humor and religion, by changing one's outlook and perspective, can precipitate anxieties into sane behavior.	B.	Humor and religion can precipitate, by changing one's outlook and perspective, anxieties into sane behavior.
12.	A.	The school board said the other residents of the neighborhood did not live in the same district as our block.	B.	The other residents of the neighborhood, the school board said, did not live in the same district as our block.

Form A/4

13. A. Men will be no more than the passive recipients of orders unless they act by the insights they have.

B. Unless they act by the insights they have, men will be no more than the passive recipients of orders.

14. A. All the people, before we had gone ten miles were crowded together on old-time open platforms.

B. All the people, were crowded together before we had gone ten miles on old-time open platforms.

15. A. As the troops raced over a rise in the land the reporters finally caught sight of the enemy.

B. The reporters, as the troops raced over a rise in the land, finally caught sight of the enemy.

16. A. Sickness and disease, by their constant yet unpredictable threat, introduce an element of tragic struggle in the world.

B. Sickness and disease introduce an element of tragic struggle in the world by their constant yet unpredictable threat.

17.	A.	Although he first studied art, in his early twenties Chesterton became a practicing journalist.	B.	In his early twenties Chesterton became a practicing journalist although he first studied art.
18.	A.	The word speed, because of its derivation from the Anglo-Saxon word sped, does not denote recklessness.	B.	Because of its derivation from the Anglo-Saxon word sped, the word speed does not denote recklessness.
19.	A.	Someone with a trained ear can distinguish, as Shaw's Professor Higgins has, the region in which a speaker resides.	B.	Someone with a trained ear, as Shaw's Professor Higgins has, can distinguish the region in which a speaker resides.
20.	A.	Human knowledge and vision are enhanced when many individual observations and thoughts are pooled.	B.	When many individual observations and thoughts are pooled, human knowledge and vision are enhanced.

Form A/6

21. A. They, having come on foot over several hundred miles, were a very tired looking people.

B. They were a very tired looking people, having come on foot over several hundred miles.

22. A. Sociologists, have rarely conducted, despite the challenges posed by higher education, insightful studies of colleges.

B. Sociologists, despite the challenges posed by higher education, have rarely conducted insightful studies of colleges.

23. A. The woman on welfare found, when the relative's will was read, that her monetary problems were over.

B. When the relative's will was read the woman on welfare found that her monetary problems were over.

24. A. Some psychologists have produced valuable information by conducting carefully controlled studies and surveys.

B. Some psychologists, by conducting carefully controlled studies and surveys, have produced valuable information.

Form A/7

25. A. I, when I went across the street to my apartment, discovered my best friend waiting impatiently.

B. I discovered, when I went across the street to my apartment, my best friend waiting impatiently.

26. A. We show ourselves to be exceedingly timid and shy by debating and agonizing over every problem.

B. By debating and agonizing over every problem, we show ourselves to be exceedingly timid and shy.

27. A. We found, watching his sleepwalking, we were most interested in where our little brother hid what he collected.

B. We, watching his sleepwalking, found we were most interested in where our little brother hid what he collected.

28. A. The troop decrease, as the evacuation rapidly commenced, was accompanied by a weakening of defenses.

B. As the evacuation rapidly commenced, the troop decrease was accompanied by a weakening of defenses.

29. A. A Scottish accent, for example has never been considered a social handicap in this country.

B. A Scottish accent, has never been considered a social handicap in this country, for example.

30. A. The vacationer becomes impressed with the complicated farm machinery while motoring through the mottled countryside.

B. While motoring through the mottled countryside, the vacationer becomes impressed with the complicated farm machinery.

31. A. The moralist has employed, while reflecting upon the brevity of human life, a variety of favorite symbols.

B. The moralist, while reflecting upon the brevity of human life, has employed a variety of favorite symbols.

32. A. After a delay that seemed to last forever, the train finally moved out of the station slowly

B. The train finally moved out of the station slowly after a delay that seemed to last forever.

APPENDIX D

APPENDIX D

SAMPLE QUESTIONNAIRE - FORM B

THIS BOOKLET CONTAINS 32 PAIRS OF SENTENCES

BOTH SENTENCES OF EACH PAIR CONTAIN
THE SAME WORDS BUT THEIR ORDER
IS DIFFERENT

PLEASE INDICATE ON THE ANSWER SHEET
WHICH SENTENCE IN EACH PAIR
YOU THINK IS HARDER

THERE ARE NO RIGHT OR WRONG ANSWERS

THERE IS NO TIME LIMIT

THANK YOU FOR YOUR HELP!
Jane E. Weber
Rutgers University

Form B/1

1. A. The train finally moved out of the station slowly after a delay that seemed to last forever.

   B. After a delay that seemed to last forever, the train finally moved out of the station slowly.

2. A. The moralist, while reflecting upon the brevity of human life, has employed a variety of favorite symbols.

   B. The moralist has employed, while reflecting upon the brevity of human life, a variety of favorite symbols.

3. A. While motoring through the mottled countryside, the vacationer becomes impressed with the complicated farm machinery.

   B. The vacationer becomes impressed with the complicated farm machinery while motoring through the mottled countryside.

4. A. A Scottish accent, has never been considered a social handicap in this country, for example.

   B. A Scottish accent, for example, has never been considered a social handicap in this country.

Form B/2

5. A. As the evacuation rapidly commenced, the troop decrease was accompanied by a weakening of defenses.

B. The troop decrease, as the evacuation rapidly commenced, was accompanied by a weakening of defenses.

6. A. We, watching his sleepwalking, found we were most interested in where our little brother hid what he collected.

B. We found, watching his sleepwalking, we were most interested in where our little brother hid what he collected.

7. A. By debating and agonizing over every problem, we show ourselves to be exceedingly timid and shy.

B. We show ourselves to be exceedingly timid and shy by debating and agonizing over every problem.

8. A. I discovered, when I went across the street to my apartment, my best friend waiting impatiently.

B. I, when I went across the street to my apartment, discovered my best friend waiting impatiently.

9. A. Some psychologists, by conducting carefully controlled studies and surveys, have produced valuable information.

B. Some psychologists have produced valuable information by conducting carefully controlled studies and surveys.

10. A. When the relative's will was read the woman on welfare found that her monetary problems were over.

B. The woman on welfare found, when the relative's will was read, that her monetary problems were over.

11. A. Sociologists, despite the challenges posed by higher education, have rarely conducted insightful studies of colleges.

B. Sociologists, have rarely conducted, despite the challenges posed by higher education, insightful studies of colleges.

12. A. They were a very tired looking people, having come on foot over several hundred miles.

B. They, having come on foot over several hundred miles, were a very tired looking people.

13.	A.	When many individual observations and thoughts are pooled, human knowledge and vision are enhanced.	B.	Human knowledge and vision are enhanced when many individual observations and thoughts are pooled.
14.	A.	Someone with a trained ear, as Shaw's Professor Higgins has, can distinguish the region in which a speaker resides.	B.	Someone with a trained ear can distinguish, as Shaw's Professor Higgins has, the region in which a speaker resides.
15.	A.	Because of its derivation from the Anglo-Saxon word *sped*, the word *speed* does not denote recklessness.	B.	The word *speed*, because of its derivation from the Anglo-Saxon word *sped*, does not denote recklessness.
16.	A.	In his early twenties Chesterton became a practicing journalist although he first studied art.	B.	Although he first studied art, in his early twenties Chesterton became a practicing journalist.

Form B/5

17. A. Sickness and disease introduce an element of tragic struggle in the world by their constant yet unpredictable threat.

B. Sickness and disease, by their constant yet unpredictable threat, introduce an element of tragic struggle in the world.

18. A. The reporters, as the troops raced over a rise in the land, finally caught sight of the enemy.

B. As the troops raced over a rise in the land the reporters finally caught sight of the enemy.

19. A. All the people were crowded together before we had gone ten miles on old-time open platforms.

B. All the people, before we had gone ten miles were crowded together on old-time open platforms.

20. A. Unless they act by the insights they have, men will be no more than the passive recipients of orders.

B. Men will be no more than the passive recipients of orders unless they act by the insights they have.

Form B/6

21. A. The other residents of the neighborhood, the school board said, did not live in the same district as our block.

B. The school board said the other residents of the neighborhood did not live in the same district as our block.

22. A. Humor and religion can precipitate, by changing one's outlook and perspective, anxieties into sane behavior.

B. Humor and religion, by changing one's outlook and perspective, can precipitate anxieties into sane behavior.

23. A. We waited breathlessly as the queen, followed by three attendants, walked gracefully out on the court.

B. We waited breathlessly as the queen walked gracefully out on the court, followed by three attendants.

24. A. Middle-Age painting embellished, by serving as the Church's handmaiden, the doctrines of Christianity.

B. Middle-Age painting, by serving as the Church's handmaiden, embellished the doctrines of Christianity.

25. A. Suppressed groups, when reaching an historical stage of awareness, grab violently for independence.

B. Suppressed groups grab violently for independence when reaching an historical stage of awareness.

26. A. Comfort is now one of the causes of its own further spread, made possible by changes in the traditional philosophy of life.

B. Made possible by changes in the traditional philosophy of life, comfort is now one of the causes of its own further spread.

27. A. Tom Dooley, skilled in medicine and fluent in French, proved valuable working in the hills of Southeast Asia.

B. Skilled in medicine and fluent in French, Tom Dooley proved valuable working in the hills of Southeast Asia.

28. A. Because of new emphasis on man and the universe, the Renaissance artist realistically depicted nature.

B. The Renaissance artist, because of new emphasis on man and the universe, realistically depicted nature.

29. A. No belief can be regarded as infallible if injustices and evils are to be eradicated.

B. No belief, if injustices and evils are to be eradicated, can be regarded as infallible.

30. A. The stranger in the corner, waving his arms in time with the music, splashed his beer all over the floor.

B. The stranger in the corner splashed his beer all over the floor, waving his arms in time with the music.

31. A. Baird, even during his daily business activities, had speculated with sending human images through space.

B. Even during his daily business activities, Baird had speculated with sending human images through space.

32. A. The people gathered around the candidate shortly after the orchestra began playing a marching song.

B. Shortly after the orchestra began playing a marching song, the people gathered around the candidate.

APPENDIX E

APPENDIX E

SAMPLE ANSWER SHEET

Form__________ Class __________

PLEASE RECORD YOUR OPINIONS HERE - CHECK THE HARDER ONE

1. _ _ A B	5. _ _ A B	9. _ _ A B	13. _ _ A B	17. _ _ A B	21. _ _ A B	25. _ _ A B	29. _ _ A B
2. _ _ A B	6. _ _ A B	10. _ _ A B	14. _ _ A B	18. _ _ A B	22. _ _ A B	26. _ _ A B	30. _ _ A B
3. _ _ A B	7. _ _ A B	11. _ _ A B	15. _ _ A B	19. _ _ A B	23. _ _ A B	27. _ _ A B	31. _ _ A B
4. _ _ A B	8. _ _ A B	12. _ _ A B	16. _ _ A B	20. _ _ A B	24. _ _ A B	28. _ _ A B	32. _ _ A B

THE KERNEL DISTANCE THEORY -
EVALUATION BY STUDENT JUDGMENT OF
SENTENCE DIFFICULTY

AN ABSTRACT OF A THESIS
SUBMITTED TO THE FACULTY
OF THE GRADUATE SCHOOL OF EDUCATION
OF
RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY
BY
JANE E. WEBER
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE
OF
MASTER OF EDUCATION

COMMITTEE CHAIRPERSON: Edward B. Fry, Ph.D.

NEW BRUNSWICK, NEW JERSEY JUNE 1977

ABSTRACT

The purpose of this study was to investigate readability of four basic sentence structure variations suggested by the Fry-Kernel Distance Theory. Hypotheses were tested with two independent samples of community college students, one with higher reading ability than the other. Subjects were given one of two versions of a questionnaire containing sentence pairs and asked to indicate on an answer sheet which sentence of each pair they judged more difficult. Patterns of decisions were evaluated for each group and then compared.

The results indicated partial support for the theory. Sentences with linguistic distance within the central kernel were judged as harder generally than those with distance at the beginning or end of a sentence. Patterns of judgments comparing splits at different places within a kernel, or imposition of distance in varying positions outside the kernel, proved to be less clear-cut. Some difference in difficulty judgment was found between the two samples, although not to the expected extent.

Possible applications for writers are mentioned along with suggestions for future research.

VITA

NAME: Jane Ellen Speyer Weber

ADDRESS: 375 Clarksville Road
Princeton Junction, New Jersey 08550

TELEPHONE: (609) 799-9060

EDUCATIONAL BACKGROUND:

High School: Jamaica High School
Jamaica, New York
Graduated in 1962

College: University of Rochester
Rochester, New York
B.A., in History Honors, 1966

University of Rochester
Rochester, New York
M.A., in Education, 1967

Clark University
Worcester, Massachusetts
Courses taken 1969-1972

PROFESSIONAL EXPERIENCE:

1966-1967: Social Studies Intern Teacher
Brighton High School
Rochester, New York

1967-Nov. 1973: Social Studies Teacher
Wachusett Regional High School
Holden, Massachusetts

1974-present: Historical Researcher
Biographical Dictionary of Princetonians
Princeton University
Princeton, New Jersey

COURSE WORK AT RUTGERS

		Instructor
Spring, 1975		
299:561	Foundations of Reading Instruction	Dr. Kling
Summer, 1975		
299:564	Remedial Reading	Dr. Zelnick
299:565	Laboratory in Remedial Reading	Dr. Zelnick
Fall, 1975		
299:608	Seminar in Reading Theories and Models	Dr. Kling
290:540	Introduction to Learning	Dr. Gilooly
610:522	Materials for Young Adults	Ms. Simpson
960:531	Statistical Methods in Education Part I	Mr. Dalal
Spring, 1976		
299:566	Seminar in Reading Research and Supervision	Dr. Kling
290:514	Introduction to Adolescent and Young Adult Years	Dr. Montare
251:530	Linguistic Bases of Language Teaching	Dr. Barone
960:532	Statistical Methods in Education Part II	Dr. Penfield
Summer, 1976		
290:509	Emotional and Social Maladjustment	Dr. Gibbons
Fall, 1976		
290:501	Introduction to Educational Tests and Measures	Dr. Geyer
299:599	Master's Thesis Research	Dr. Fry